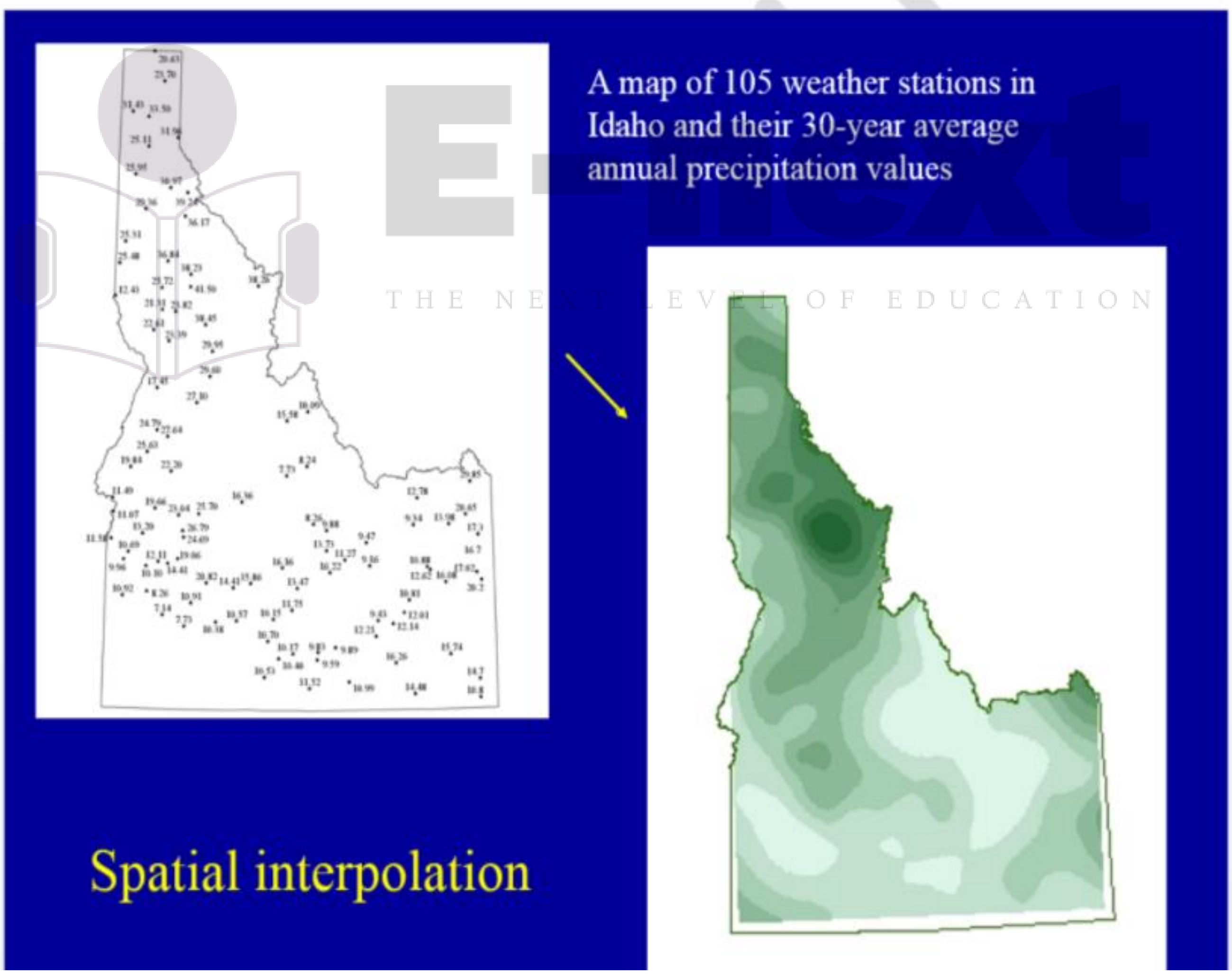
UNIT –VI

Spatial Interpolation: Elements, Global methods, local methods, Kriging, Comparisons of different methods

Q.1) What is Spatial Interpolation?

- Spatial interpolation is the process of using points with known values to estimate values at other points.
- These points with known values are called known points, control points, sampled points, or observations.
- In GIS applications, spatial interpolation is typically applied to a grid with estimates made for all cells.
- Spatial interpolation is therefore a means of converting point data to surface data so that the surface data can be used with other surfaces for analysis and modeling.



Q.2) Elements of Spatial Interpolation & its types.

Elements of Spatial Interpolation

- Spatial interpolation requires two basic points and an interpolation method.

Control Points

- Control points are points with known values.
- Control points provide the data necessary for the development of interpolator for spatial interpolation.
- Two factors can influence the accuracy of Spatial Interpolation
    1) Number of control points
    2) Distribution of control points.
- A basic assumption in spatial interpolation is that the value to be estimated at a point is more influenced by nearby known points than the points farther away.
- To be effective for estimation the points should be well distributed within the study area.

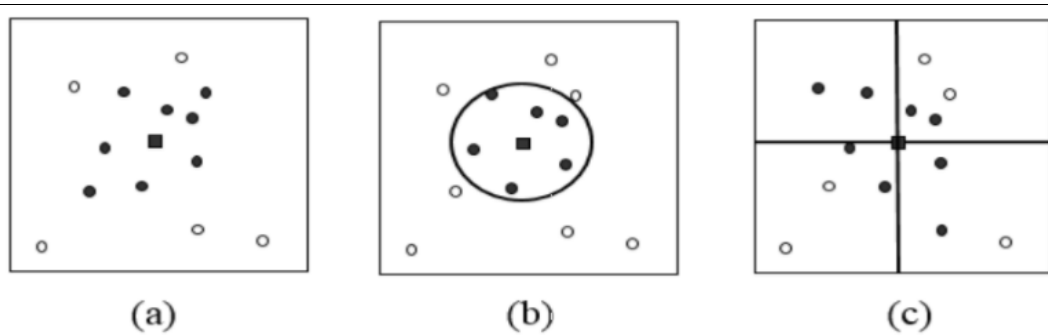
Type of Spatial Interpolation

- Spatial interpolation methods can be categorized in several ways.

1) Global & Local interpolation
2) Exact interpolation & Inexact Interpolation
3) Deterministic and Stochastic

1) Global & Local interpolation

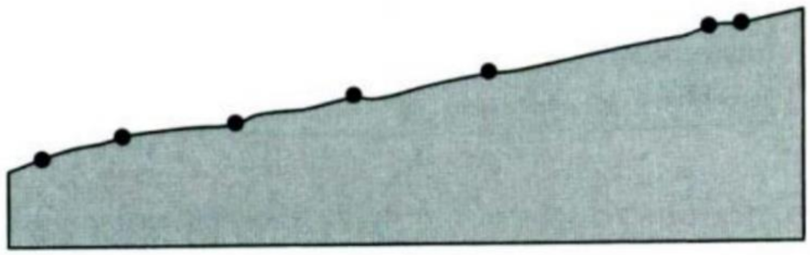
- A global interpolation method uses every known point available to estimate an unknown value.
- A local interpolation method, on the other hand, uses a sample of known points to estimate an unknown value.



Three search methods for sample points:
(a) find the closest points to the point to be estimated,
(b) find points within a radius, and
(c) find points within each of the four quadrants.

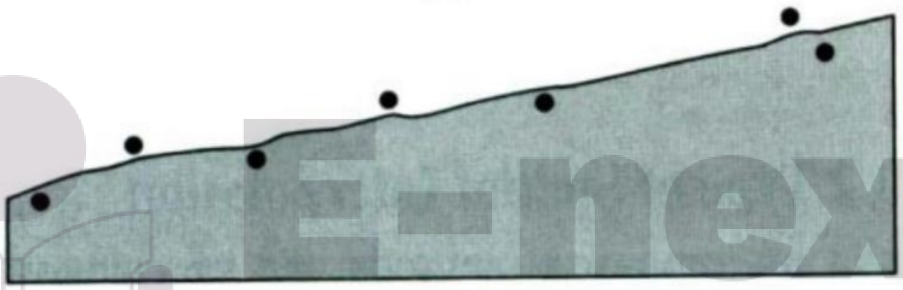
2) Exact vs. Inexact

- Exact interpolation predicts a value at the point location that is the same as its known value.
- In other words, exact interpolation generates a surface that passes through the control points.



Exact Interpolation

- In contrast, inexact interpolation or approximate interpolation predicts a value at the point location that differs from its known value.



Inexact Interpolation

Deterministic vs. Stochastic

- A deterministic interpolation method provides no assessment of errors with predicted values.
- A stochastic interpolation method, on the other hand, offers assessment of prediction errors with estimated variances.

Q.3 Explain Global methods, local methods, Kriging?

A Classification of Spatial Interpolation Methods

Global		Local	
Deterministic	Stochastic	Deterministic	Stochastic
Trend surface (inexact)*	Regression (inexact)	Thiessen (exact)	Kriging (exact)
		Density estimation (inexact)	
		Inverse distance weighted (exact)	
		Splines (exact)	

Global Methods

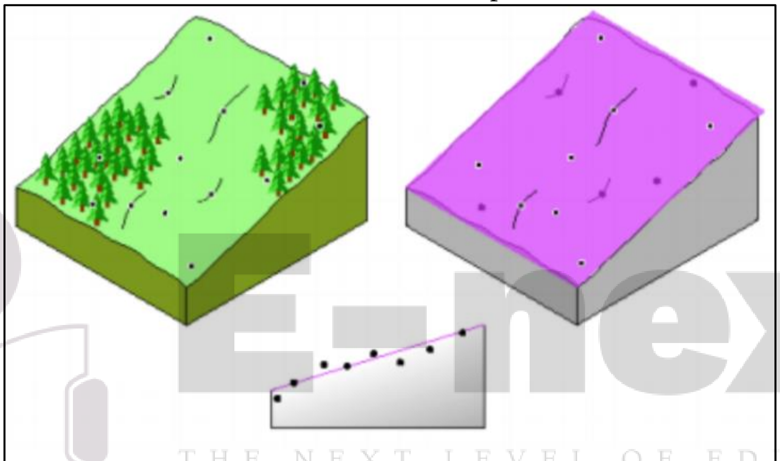
Trend Surface Analysis

- The inexact interpolation method, Trend Surface Analysis approximates points with known values with a polynomial equation.
- The equation or interpolator can then be used to estimate the values at other points.
- A linear First order Trend Surface Model uses the equation

First-order trend surface (polynomial)

$$z_{x,y} = b_0 + b_1 x + b_2 y$$

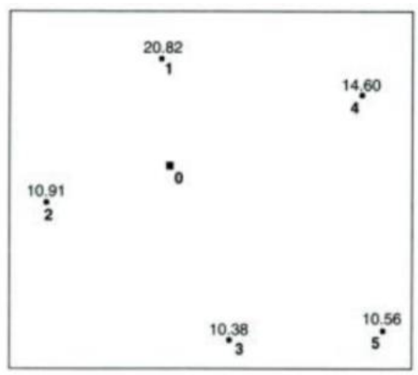
- Where the attribute value z is the function of x and y coordinates
- The b coefficeints are estimated from the known points



First Order Trend Surface

- A worked example

Point	x	y	Value
1	69	76	20.820
2	59	64	10.910
3	75	52	10.380
4	86	73	14.600
5	88	53	10.560
0	69	67	?



(1) Set up 3 equations

$$\sum z = b_0 n + b_1 \sum x + b_2 \sum y$$
$$\sum xz = b_0 \sum x + b_1 \sum x^2 + b_2 \sum xy$$
$$\sum yz = b_0 \sum y + b_1 \sum xy + b_2 \sum y^2$$

(2) Re-write in matrix format

$$\begin{bmatrix} n & \sum x & \sum y \\ \sum x & \sum x^2 & \sum xy \\ \sum y & \sum xy & \sum y^2 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum z \\ \sum xz \\ \sum yz \end{bmatrix}$$

(3) Calculate

$\sum X = 377$
$\sum Y = 318$
$\sum X^2 = 29007$
$\sum Y^2 = 20714$
$\sum XY = 23862$
$\sum YZ = 4445.8$
$\sum XZ = 5044$

X	Y	Z	X²	Y²	XY	XZ	YZ
69	76	20.82	4761	5776	5244	1437	1582.32
59	64	10.91	3481	4096	3776	643.7	698.24
75	52	10.38	5625	2704	3900	778.5	539.76
86	73	14.6	7396	5329	6278	1256	1065.8
88	53	10.56	7744	2809	4664	929.3	559.68
377	**318**	**67.27**	**29007**	**20714**	**23862**	**5044**	**4445.8**

(4) Plug in values for 5 points

$$\begin{bmatrix} 5 & 377 & 318 \\ 377 & 29007 & 23862 \\ 318 & 23862 & 20714 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 67.270 \\ 5043.650 \\ 4445.800 \end{bmatrix}$$

(5) Solve for b coefficients:
Multiply inverse of left matrix by right matrix

$$\begin{bmatrix} 23.2102 & -0.1631 & -0.1684 \\ -0.1631 & 0.0018 & 0.0004 \\ -0.1684 & 0.0004 & 0.0021 \end{bmatrix} \cdot \begin{bmatrix} 67.270 \\ 5043.650 \\ 4445.800 \end{bmatrix} = \begin{bmatrix} -10.094 \\ 0.020 \\ 0.347 \end{bmatrix}$$

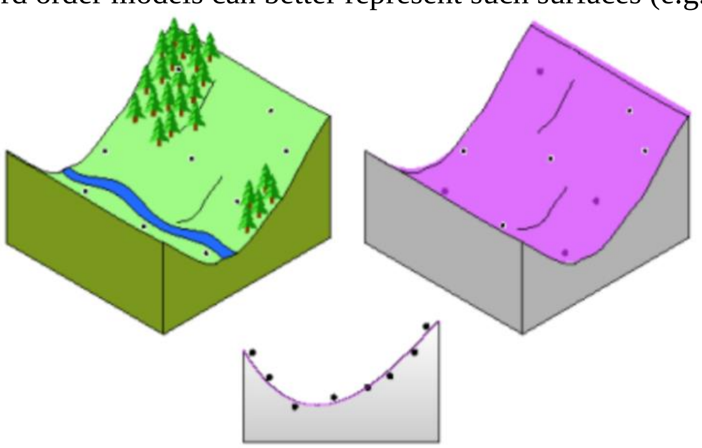
(6) Use the b coefficients to calculate "z" for any point (X,Y) (69, 67)

$$z_{x,y} = b_0 + b_1 x + b_2 y$$

$$z_0 = -10.094 + (0.020)(69) + (0.347)(67) = 14.535$$

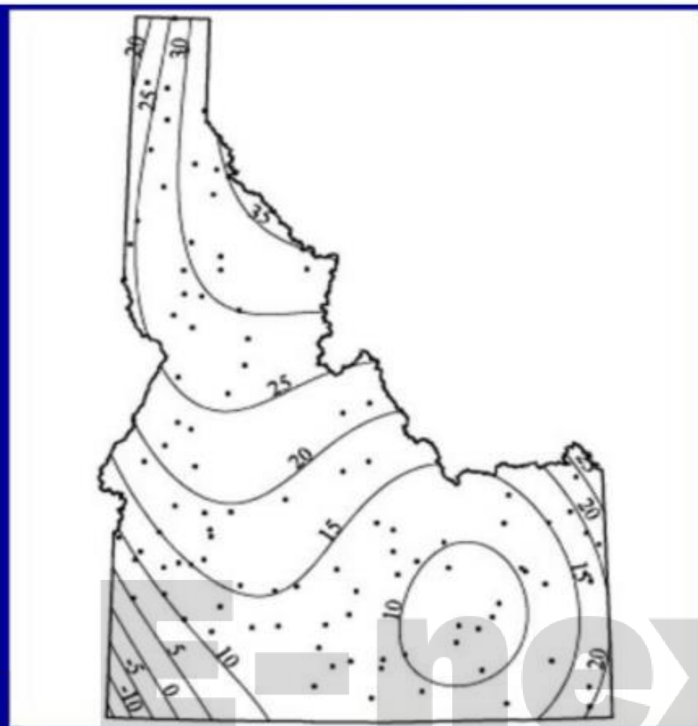
Higher-order trend surface

- First order polynomials (inclined surface) can not represent the complex natural surfaces.
- A cubic or third order models can better represent such surfaces (e.g., hills, valleys)



Third-order trend surface
(nine coefficients)

$$z_{x,y} = b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 + b_6x^3 + b_7x^2y + b_8xy^2 + b_9y^3$$



An isoline map of a third-order trend surface created from 105 points with annual precipitation values

Regression Models

- Regression model relates a dependent variable to a number of independent variables in an equation which can then be used for prediction or estimation.
- Non-spatial models should not be used.

Local Method

- It is all about mechanisms for the selection of a sample of control points.

6.4.1 Thiessen Polygons
6.4.2 Density Estimation
6.4.3 Inverse Distance Weighted Interpolated
6.4.4 Thin-plate Splines
6.4.5 Kriging
6.4.5.1 Ordinary Kriging
6.4.5.2 Universal Kriging
6.4.5.3 Other Kriging Methods

6.4.1 Thiessen Polygons:

- Thiessen polygons assume that nay point within a polygon is closer to the polygon's known point than any other known point.
- Thiessen polygons were originally proposed to estimate areal averages of precipitation by making sure that any point within a polygon is closer to the polygon's weather station than any other station.
- Thiessen polygons, also called Voronoi polygons, are used in a variety of applications, especially for service area analysis of public facilities such as hospitals.
- Thiessen polygons do not use an interpolator but require initial triangulation for connecting known points.
- Because different ways of connecting points can form different sets of triangles, the Delaunay triangulation triangulation ensures that each known point is connected to its nearest neighbors, and that triangles are as equilateral as possible.
- After triangulation, Thiessen polygons can be easily constructed by connecting lines drawn perpendicular to the sides of each triangle at their midpoints (Figure).

Figure 1
Thiessen polygons (in thicker lines) are interpolated from the known points and the Delaunay triangulation (in thinner lines).

- Thiessen polygons are smaller in areas where points are closer together and larger in areas where points are farther apart.
- This size differentiation is the basis, for example, for determining the quality of public service.
- A large polygon means greater distances between home locations and public service providers.
- The size differentiation used for other purposes such as predicting forest age classes, with larger polygons belonging to older trees .

6.4.2 Density Estimation

- Density estimation measures cell densities in a raster by using a sample of known points.
- There are simple and kernel density estimation methods.

- To use the simple density estimation method, we can place a raster on a point distribution, tabulate points that fall within each cell, sum the point values, and estimate the cell's density by dividing the total point value by the cell size.
- Figure 16.7 shows the input and output of an example or simple density estimation.
- The input is a distribution of sighted deer locations plotted with a 50-meter interval to accommodate the resolution of telemetry.
- A circle, rectangle, wedge, or ring based at the center of a cell may replace the cell in the calculation.
- Kernel density estimation associates each known point with a kernel function for the purpose of estimation .
- Expressed as a bivariate probability density function a kernel function looks like a bump, centering at a Known point and tampering off to 0 over a defined bandwidth or window area.
- The kernel function and the bandwidth determine the shape of the bump which in turn determines the amount of smoothing in estimation.

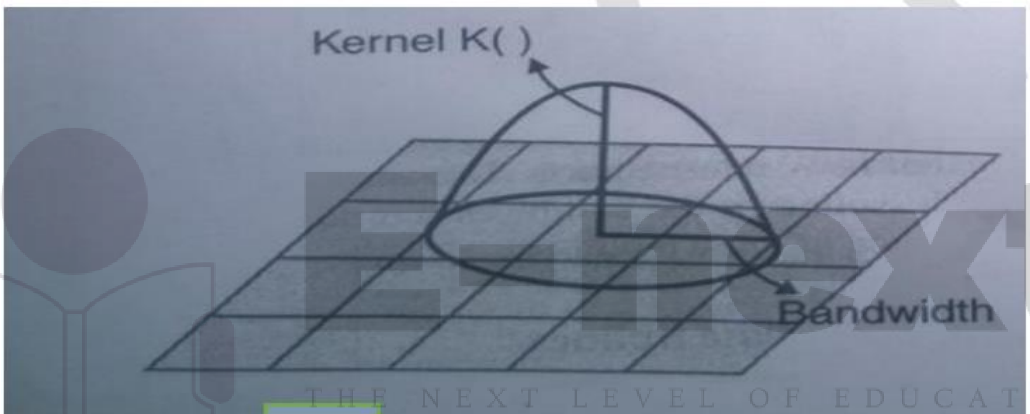


Figure
A kernel function, which represents a probability density function, looks like a "bump" above a grid.

- The kernel density estimator at point x is then the sum of bumps placed at the known points x, within the bandwidth:

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{1}{h}(x - x_i)\right)$$

- where K( ) is the kernel function, h is the band-width, n is the number of known points within the bandwidth and d is the data dimensionality.
- For two-dimensional data (d= 2), the kernel Function is usually given by:

6.4.3 Inverse Distance weighted interpolation

- Inverse distance weighted (IDW) interpolation, is an exact method that enforces the condition that the estimated value of a point is influenced more by nearby known points than by those farther away.

 The general equation for the IDW method is:

z_0	Estimated value at point 0
z_i	Is the z value at control point i
d_i	Distance between point I and point 0
k	The larger the k, the greater the influence of neighboring points.
S	number of used points

$$z_0 = \frac{\sum_{i=1}^{s} z_i \frac{1}{d_i^k}}{\sum_{i=1}^{s} \frac{1}{d_i^k}}$$

Z_i	d_i	d_i^2	$1/(d_i^2)$	$Z_i \times 1/(d_i^2)$
20.82	18	324	0.0031	0.06426
10.91	20.88	435.97	0.0023	0.02502
10.38	32.31	1043.9	0.0010	0.00994
14.6	36.05	1299.6	0.0008	0.01123
10.56	47.2	2227.8	0.0004	0.00474
SUM			0.0076	0.11520

z_i	Between points	Distance (d_i)
20.82	0,1	18
10.91	0,2	20.88
10.38	0,3	32.31
14.6	0,4	36.05
10.56	0,5	47.20

Example

Assuming k = 2

$$\sum z_i\, 1/d_i^2 = (20.820)(1/18.000)^2 + (10.910)(1/20.880)^2 + (10.380)(1/32.310)^2 + (14.600)(1/36.056)^2 + (10.560)(1/47.202)^2 = 0.1152$$

$$\sum 1/d_i^2 = (1/18.000)^2 + (1/20.880)^2 + (1/32.310)^2 + (1/36.056)^2 + (1/47.202)^2 = 0.0076$$

$$z_0 = 0.1152/0.0076 = 15.158$$

6.4.4 Thin-Plate Splines

 Splines for spatial interpolation are conceptually similar to splines for line smoothing except that in spatial interpolation they apply to surfaces rather than lines.
 Thin-plate splines create a surface that passes through the control points and has the least possible change in slope at all points.
 In other words, thin-plate splines fit the control points with a minimum curvature surface.
 The approximation of thin-plate splines is of the form:

$$Q(x, y) = \sum A_i d_i^2 \log d_i + a + bx + cy$$

 -where x and y are the x-, y-coordinates of the point to be interpolated,

$$d_i^2 = (x - x_i)^2 + (y - y_i)^2,$$

 and xi and yi are the x-, y-coordinates of control point i.
 Thin-plate splines consist of two components: (a + bx + cy) represents the local trend function, which has the same form as a linear or first-order-trend surrace.
 and di^2 log d represents a basis function, which is designed to obtain minimum curvature surfaces.
 The coefficients Ai and a, b, and c are determined by a linear system of –equations

$$\sum_{i=1}^{n} A_i d_i^2 \log d_i + a + bx + cy = f_i$$

$$\sum_{i=1}^{n} A_i = 0$$

$$\sum_{i=1}^{n} A_i x_i = 0$$

$$\sum_{i=1}^{n} A_i y_i = 0$$

 Where n is the number of control points, and fi is the known value at control point i.
 The estimation of the coefficients requires n + 3 simultaneous equations.

6.4.5 Kriging

 Kriging is a stochastic model that provides estimates for accuracy/certainty in predictions.
 Kriging is a geostatistical method for spatial interpolation.
 A statistical based estimator of spatial variables.
 Kriging differs from interpolation methods discussed so far because kriging can assess the quality of prediction with estimated prediction errors.
 Assumes spatial variation is neither totally random nor deterministic.
 Components:

 Spatial trend (an increase/decrease in a variable that depends on direction e.g. temperature may decrease toward the northwest)
 Autocorrelation (the tendency for points near each other to have similar values)
 Random (stochastic)
 Creates a mathematical model which is used to estimate values across the surface
 Presence or absence of a drift and the interpretation of the regionalized variable have led to development of different Kriging methods.

o Ordinary Kriging: assumes absence of drift, focuses on spatially correlated component
o Universal Kriging assumes that the spatial variation in z values has a drift or variation in addition to the spatial correlation.
o Other: block Kriging, co-Kriging.

Semivariogram

- Kriging uses the semivariance to measure the spatially correlated component, a component that is also called spatial dependence or spatial correlation.
- Semivariograms measure the strength of statistical correlation as a function of distance; they quantify spatial autocorrelation.
- Because Kriging is based on the semivariogram, it is probabilistic, while IDW and Spline are deterministic.
- Kriging associates some probability with each prediction, hence it provides not just a surface, but some measure of the accuracy of that surface.
- The semivariance is computed by

$$\gamma(h) = \frac{1}{2}[z(x_i) - z(x_j)]^2$$

Where ,

$\gamma(h)$ is the semivariance between known points

x_i and x_j separated by the distance h

z is the attribute value.

 If spatial dependence does exist in a data set , known points that are close to each other are expected to have small semivariance, and known points that are farther apart are expected to have larger semivariance.
 A process called binning is typically used in Kriging to average semivariance data by distance and direction.
 The first part of the binning process is to group pairs of sample points into lag classes.
 For example , if the lag size is 2000 meters , then pairs of points separated by less than 2000 meters are grouped into lag class of 0-2000, pairs of points separated between 2000 and 4000metres are grouped into the lag class of 2000-4000, and so on.
 The second part of the binning process is to group pairs of sample points by direction.
 The result of the binning process is a set of bins that sort pairs of sample points by distance and direction.
 The next step is to compute the average semivariance by:

$$\gamma(h) = \frac{1}{2n}\sum_{i=1}^{n}[z(xi) - z(xi + h)]$$

 γ(h) is the average semivariance between known points separated by lag h;
 n is the number of pairs of sample points sorted by direction in the bin
 z is the attribute value
 A semivariogram plots the average semivariance against the average distance.\
 Because of the directional component, one or more average semivarience may be plotted at the same distance.
 We can examine the semivariogram in figure A by distance.

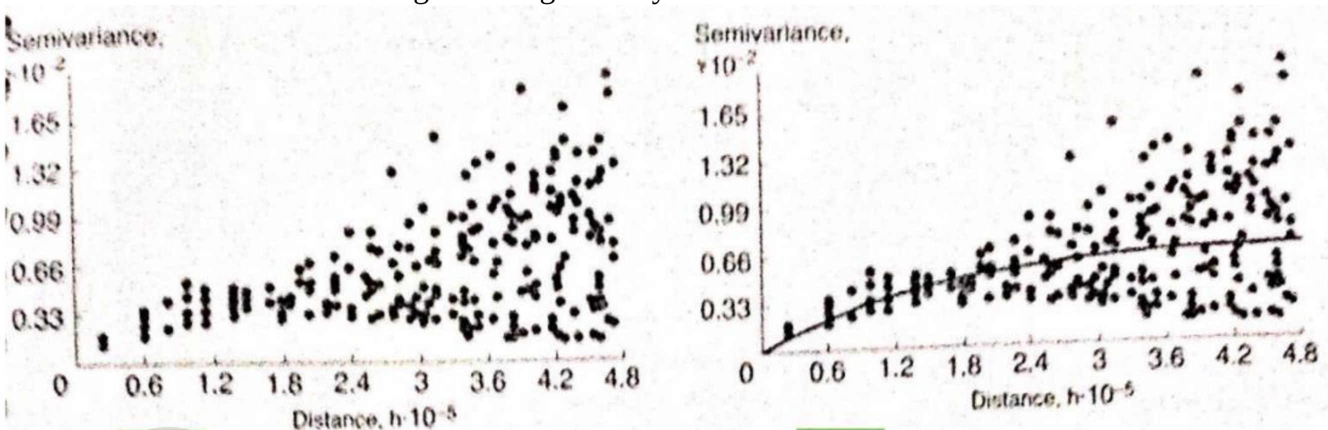


Figure A
A semivariogram after binning.

Figure B
Fitting a semivariogram with a mathematical function or a model.

 If Spatial dependence exists among the sample points then pairs of points that are closer in distance will have more similar values than pairs that are farther apart.
 In other words the semivariance is expected to increase as their distance increases in the presence of the spatial dependence.
 Are semivariogram can also be examined by direction.
 If spatial dependence has directional differences then the semivariance values may change more rapidly in one direction than another.
 Anisotropy is the term describing the existence of directional differences in spatial dependence.

Models:-

 A semivariogram such as figure A may be used alone as a measure of spatial autocorrelation in the data set.
 But to be used as an interpolator in kriging , the semivariogram must be mathematical function or model (figure B)
 Two common models for fitting semivariogram are spherical and exponential.
 A spherical model shows a progressive decrease of spatial dependence until some distance , beyond which the spatial dependence levels off.
 In exponential model spatial dependence decreases exponentially with increasing distance and disappears completely at an infinite distance.

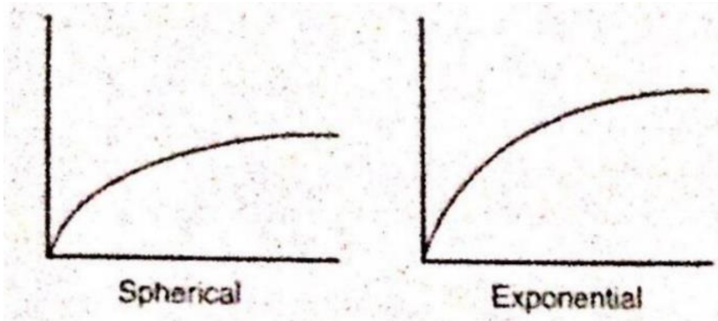
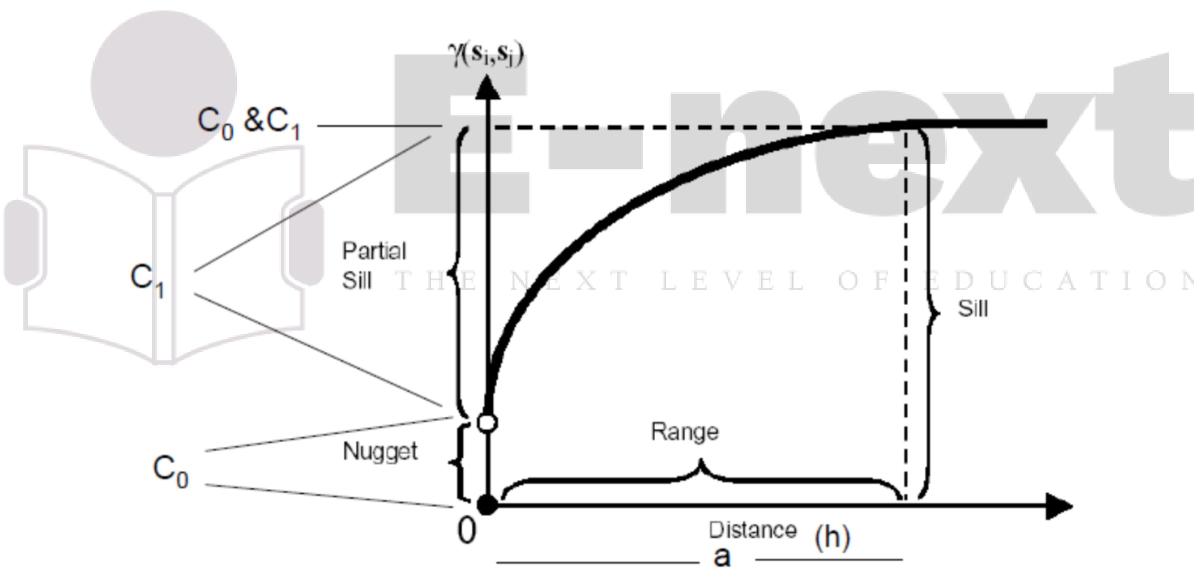


Figure C
Two common models for fitting semivariograms: spherical and exponential.

6.4.5.1 Ordinary Kriging
6.4.5.2 Universal Kriging
6.4.5.3 Other Kriging Methods

- A fitted semivariogram can be dissected into three possible elements : nugget, range, and sill (figure)



- The nugget is the semivariance at the distance of zero, representing measurement error or microscale variation, or both.
- The range is the distance at which the semivariance starts to level off.
- In other words, the range corresponds to the spatially correlated portion of the semivariogram beyond the range, the semivariance becomes a relatively constant value.
- The semivariance at which the leveling takes place is called the sill.
- The sill comprises two components: the nugget and the partial sill.
- To put it another way, the partial sill is the difference between the sill and the nugget.

6.4.5.1 Ordinary Kriging

- Assuming the absence of a drift, ordinary kriging focuses on the spatially correlated component and uses the fitted semivariogram directly for interpolation.

- The general equation for estimating the z value at a point is

$$z_0 = \sum_{i=1}^{s} z_x W_x$$

where z_0 is the estimated value, z_x is the known value at point x, W_x is the weight associated with point x, and s is the number of sample points used in estimation. The weights can be derived from solving a set of simultaneous equations. For example, the following equations are needed for a point (0) to be estimated from three known points (1, 2, 3):

$$W_1\gamma(h_{11}) + W_2\gamma(h_{12}) + W_3\gamma(h_{13}) + \lambda = \gamma(h_{10})$$
$$W_1\gamma(h_{21}) + W_2\gamma(h_{22}) + W_3\gamma(h_{23}) + \lambda = \gamma(h_{20})$$
$$W_1\gamma(h_{31}) + W_2\gamma(h_{32}) + W_3\gamma(h_{33}) + \lambda = \gamma(h_{30})$$
$$W_1 + W_2 + W_3 + 0 = 1.0$$

where $\gamma(h_{ij})$ is the semivariance between known points i and j, $\gamma(h_{i0})$ is the semivariance between the ith known point and the point to be estimated, and λ is a Lagrange multiplier, which is added to ensure the minimum possible estimation error. Once the weights are solved, Eq. (16.14) can be used to estimate z_0

$$z0 = z1W1+z2W2+z3W3$$

- The preceding example shows that weights used in kriging involve not only the semivariances between the point to be estimated and the known points but also those between the known points.
- This differs from the IDW method, which uses only weights applicable to the point to be estimated and the known points.
- Another important difference between kriging and other local methods is that kriging produces a variance measure for each estimated point to indicate the reliability of the estimation.

6.4.5.2 Universal Kriging

- Universal kriging assumes that the spatial variation in z values has a drift or a trend in addition to the spatial correlation between the sample points.
- Typically , universal kriging incorporates a first order or the second order polynomial in the kriging process.
- A first order polynomial is

$$M=b1xi+b2y$$

Where M is the drift

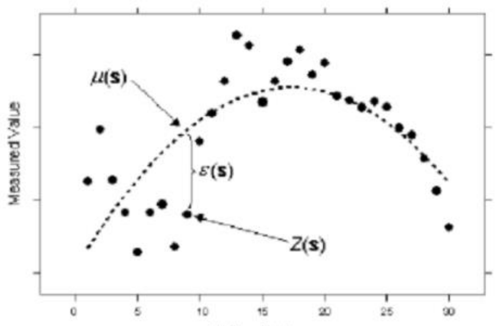
Xi and yi are the x and y coordinates of the sampled point i, and b1 and b2 are the drift coefficients

- A second order polynomial is

$$M=b_1x_i+b_2y_i+b_3x_i^2+b_4x_iy_i+b_5y_i^2$$

Other kriging methods:-

Components of Kriging



The value of z depends on: (1) trend component, (2) random autocorrelated component, and (3) random non-correlated component (for simplicity not represented in figures)

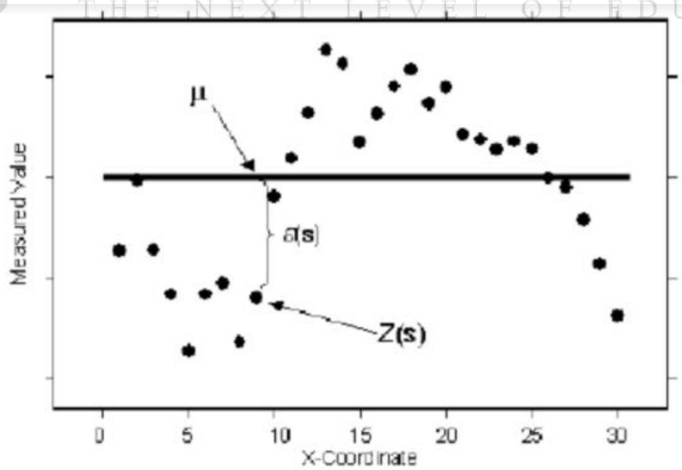
$$Z(s) = \mu(s) + \varepsilon(s)$$

Where:

Z = Value at point s
μ = Trend component value at point s (first order or second order polynomial)
ε = Random, autocorrolated component

Simple kriging

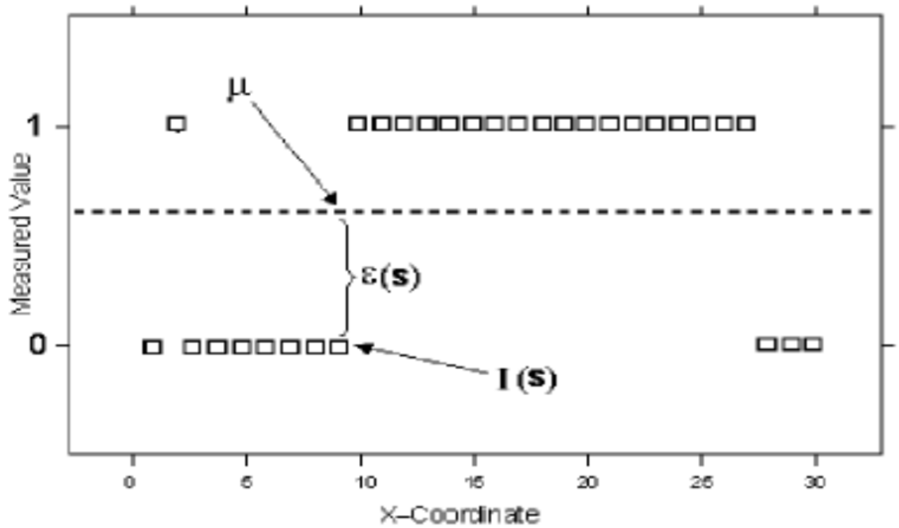
 Assumes that the mean of the data set is known.



- Assumes μ(s) , the mean of data set is known and is constant
- Assumes there is no trend component
- In the majority of cases this is unrealistic assumption

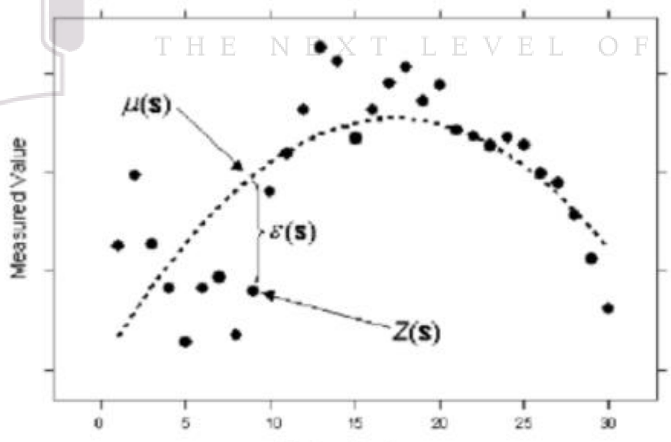
Indicator Kriging

- uses binary data rather than continuous data.



- μ(s) is constant, and unknown
- Values are binary (1 or 0)
- Example, a point is forest or non forest

Universal Kriging



- Assumes z values change because of a drift (trend) in addition to autocorrelation.
- μ(s) is not constant
- Trend component expressed as a 1st order (plane) or 2nd order polynomials (quadratic surface)
- Kriging is performed on residual after the trend is removed